International Journal of Research and Development in Engineering Sciences
ISSN: 2582 - 4201(Online) , http://www.ijrdes.com/
Regular Issue, Vol. 6, Issue. 3 (May – June) , 2024 , Pages: 34 - 37
Received : 20 April 2024 ; *Accepted*: 28 April 2024 ; *Published*: 01 June 2024
https://doi.org/10.63328/IJRDES-V6RI3P6
Original : Applied Research – Experimental Research Paper

Detection of Chronic Kidney Disease using Machine Learning Algorithms

R. Triveni [1 *], M. Kalpana [2], T Indiramma [3], B P Deepika [4], S Bhargavi [5]

[1-5] Department of Electronics and Communication Engineering, Jawaharlal Nehru University, Anantapuramu, Andhra Pradesh , India;

** Corresponding Author: R. Triveni ; triveni.aishu@gmail.com*

Abstract: The diagnosis of kidney disease, often called chronic renal illness, is known as chronic kidney disease (CKD). Chronic kidney disease (CKD) is a widespread and chronic health issue that requires preventative measures for early identification in order to slow the disease's progression. Using complex techniques like Support Vector Machine (SVM) and Logistic Regression, this study explores the field of machine learning. For robust model construction, use Random Forest and Decision Tree. Through the use of Variance Inflation Factor (VIF) for feature engineering, the dataset is carefully refined to improve the visibility of relevant features. To address any possible data imbalance concerns, Synthetic Minority Over-sampling Technique (SMOTE) is also utilized, promoting a fairer representation of classes in the dataset. By means of a thorough assessment procedure, the study methodically pinpoints the key elements that contribute to a precise diagnosis of chronic kidney disease. Each algorithm's effectiveness is evaluated using performance indicators like recall, accuracy, precision, and Fl-score.

Keywords: Learning, Chronic Kidney Disease(CKD), Dataset, Support Vector Machine(SVM), Decision Tree, ANN.

1. Introduction

Chronic kidney disease (CKD), which is defined by a progressive loss of kidney function over time, is a major global public health concern. Its growing prevalence and morbidity rates are attributed to a number of risk factors, including cardiovascular disease, diabetes mellitus, and hypertension. Accurate diagnostic methods and predictive models are crucial for mitigating the progression of chronic kidney disease (CKD) and its accompanying consequences. Early detection and intervention are essential for this. The promise of machine learning (ML) and data-driven techniques to enhance the diagnosis and prognosis of chronic kidney disease (CKD) has drawn interest.

Personalized treatment plans, risk profile stratification, and patient identification of individuals with chronic kidney disease (CKD) can all be made more accurate and efficient with the use of machine learning (ML) techniques that make use of large-scale datasets and complex algorithms. This convergence of technology and healthcare offers a rare chance to transform the treatment of chronic kidney disease and improve patient outcomes.

1.1. Image Segmentation In CKD Diagnosis

The process of diagnosing chronic kidney disease (CKD) by image segmentation entails using sophisticated imaging methods to precisely identify anomalies and kidney structures. For patients with chronic kidney disease (CKD), this procedure is essential for evaluation, tracking, and treatment planning. The use of image segmentation in the diagnosis of chronic kidney disease is described in full here. When it comes to diagnosing chronic kidney disease (CKD), image segmentation can be quite helpful, particularly when it comes to MRI and CT scans. Clinicians can examine particular abnormalities or characteristics, such as cysts, tumors, or structural alterations, that are suggestive of chronic kidney disease (CKD) by segmenting the kidney region from these pictures.

1.2. Multi-Omic Insights For Targeted CKD Therapy

Through the integration of Multi-Omic data and the execution of extensive molecular profiling studies, the research seeks to clarify the pathophysiological mechanisms behind the progression and consequences of chronic kidney disease.

ISSN: 2582 - 4201(Online) , http://www.ijrdes.com/, IJRDES, Vol. 6, Issue 3, 2024 , https://doi.org/10.63328/IJRDES-V6RI3P6

In order to find novel therapeutic targets and biomarkers for early detection, prognosis, and targeted therapy, the research aims to untangle the intricate interplay between genetic, epigenetic, and environmental variables leading to the pathogenesis of CKD. In order to provide integrated treatment plans that meet the comprehensive needs of CKD patients, the research also attempts to improve our knowledge of the relationships between comorbid diseases like hypertension, diabetes mellitus, and cardiovascular disease and chronic kidney disease.

2. Related Work

Anupama et al. created a model for diagnosing intracerebral haemorrhage(ICH) based on the Grab Cut segmentation technique. The model used Gabor filtering to remove noise and improve picture quality, and synergic deep learning (SDL) for feature extraction [1].Using the stacked autoencoder model to extract the most valuable and effective features from the chronic renal disease dataset and the SoftMax classifier for prediction, Gupta et al. developed a deep learning classification framework that produced high accuracy predictions [2].

Rodrigues et al. developed a hybrid model that combines Recurrent Neural Networks (RNN) and Convolutional Neural Networks (CNN) to offer visually descriptive captions and high-definition visuals for videos. The features of the video pictures were extracted by CNN, and the Long Short-Term Memory (LSTM) was trained with the SoftMax function to produce phrases with a meaningful sequence [3].

Using 50-layer Residual Neural Networks (ResNet) for the diagnosis and identification of different skin illnesses, Sharma et al. created an expert system to assist skin doctors [4]. Radiologists can use Wang et al.'s suggested AlexNet model for alcoholism identification to diagnose patients. The primary transfer learning model, AlexNet, was employed together with five data augmentation (DA) approaches to enhance classification performance: random translation, gamma correction, noise injection, scaling, and image rotation.

3. Experimental Procedure

Establish which variables are dependent and independent. A sample size that has adequate statistical power to identify differences between groups should be chosen. Determine how long the study will last, keeping in mind that CKD is a chronic condition.

3.1. Data Collection

The chronic kidney disease detection project gathers pertinent medical datasets, including patient information, test results, and diagnostic records, during the data gathering phase. Important characteristics including blood pressure, age, serum creatinine levels, and other relevant clinical indications are the main focus of data gathering.

3.2. Data Preprocessing

This include dealing with outliers, handling missing numbers, and scaling or normalizing numerical features. The Variance Inflation Factor (VIF) and other feature engineering approaches are used to improve the discriminatory power of the dataset by encoding categorical variables.

3.3. Model Training and Evaluation

To properly train the models, the pre-processed data is divided into training and testing sets. In order to accurately detect chronic kidney disease, the algorithms optimize their parameters by identifying patterns and linkages within the data.

An extensive evaluation is conducted on the trained models' performance using a number of metrics, such as accuracy, precision, recall, and F1-score. The models' capacity to generalize to previously unseen data is evaluated using a different test dataset.

4. Results and Discussion

Tensorflow and Python 3.8.0 are used to implement the suggested algorithms-based approach in our numerical results.

4.1. Support Vector Machine

```
svm1=SVC()
svm1.fit(X_train,y_train)
pred_svm1=svm1.predict(X_test)
print(accuracy score(pred svm1,y test))
print(classification_report(pred_svm1,y_test))
✓ 0.0s

0.6736842105263158
              precision    recall  f1-score   support

           0       0.74      0.65      0.69        54
           1       0.60      0.71      0.65        41

    accuracy                           0.67        95
   macro avg       0.67      0.68      0.67        95
weighted avg       0.68      0.67      0.68        95
```

Fig.1 Classification Report for Support Vector Machine

ISSN: 2582 - 4201(Online) , http://www.ijrdes.com/, IJRDES, Vol. 6, Issue 3, 2024 , https://doi.org/10.63328/IJRDES-V6RI3P6

4.2. Logistic Regression

```
lr1=LogisticRegression()
lr1.fit(X_train,y_train)
pred_lr1=lr1.predict(X_test)
print(accuracy_score(pred_lr1,y_test))
print(classification_report(pred_lr1,y_test))
```

```
0.9789473684210527
              precision    recall  f1-score   support

           0       0.96      1.00      0.98        45
           1       1.00      0.96      0.98        50

    accuracy                           0.98        95
   macro avg       0.98      0.98      0.98        95
weighted avg       0.98      0.98      0.98        95
```

Fig.2 Classification Report for Logistic Regression

4.3. Decision Tree

```
DT1=DecisionTreeClassifier()
DT1.fit(X_train,y_train)
pred_DT1=DT1.predict(X_test)
print(accuracy_score(pred_DT1,y_test))
print(classification_report(pred_DT1,y_test))
```

```
0.9894736842105263
              precision    recall  f1-score   support

           0       1.00      0.98      0.99        48
           1       0.98      1.00      0.99        47

    accuracy                           0.99        95
   macro avg       0.99      0.99      0.99        95
weighted avg       0.99      0.99      0.99        95
```

Fig.3 Classification Report for Decision Tree

4.4. Random Forest

```
rf1=RandomForestClassifier()
rf1.fit(X_train,y_train)
pred_rf1=rf1.predict(X_test)
print(accuracy_score(pred_rf1,y_test))
print(classification_report(pred_rf1,y_test))
```

```
1.0
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        47
           1       1.00      1.00      1.00        48

    accuracy                           1.00        95
   macro avg       1.00      1.00      1.00        95
weighted avg       1.00      1.00      1.00        95
```

Fig: 4 Classification Report for Random Forest

4.5. ANN Classifier

```
loss, accuracy = neural_network.evaluate(X_test, y_test)
print(f'Test Loss: {loss}')
print(f'Test Accuracy: {accuracy}')
```

```
3/3 ━━━━━━━━━━━━━━━━━━━━ 0s 13ms/step - accuracy: 0.7110 - loss: 3.4749
Test Loss: 2.9283287525177
Test Accuracy: 0.7578947544097 9
```

Fig. 5 Classification Report for ANN

Using SVM, Logistic Regression, Decision Tree, Random Forest, and ANN classifiers, we obtain accuracy of 67%, 98%, 99%, 100%, and 75% based on the input data, respectively. The following provides a detailed comparison of the five algorithms: SVM, ANN Classifier, Random Forest, Decision Tree, and Logistic Regression.

Equation/Formula

$$y = \frac{e^{(b_0 + b_1 X)}}{1 + e^{(b_0 + b_1 X)}}$$

$$\text{Accuracy} = \frac{TP+NP}{TP+FP+TN+FN}$$

CLASSIFERS	ACCURACY %	PRECISION		RECALL		F1 SCORE		SUPPORT	
		0	1	0	1	0	1	0	1
Logistic Regression	88	88	89	89	88	88	88	47	48
SVM	51	33	68	52	50	41	58	31	64
Decision Tree	98	96	100	100	96	98	98	46	49
Random Forest	98	98	100	100	98	99	99	47	48
ANN	91	86	87	97	98	91	92	45	44

Fig. 6 Comparison between five Algorithms

5. Conclusion

Machine learning techniques such as logistic regression, support vector machines, Random Forests, Decision Trees, and ANN classifiers are employed in the diagnosis of chronic renal disease. Our accuracy was approximately 67%, 98%, 99%, 100%, and 75%, in that order. It is preferable to use a functional product in future research to identify chronic renal illness.

Declaration

To the best of our knowledge, we certify that the research for this article is entirely original, with no information from outside sources utilized in its composition. We solely bear responsibility for any legal and copyright issues if such sort material is uncovered in subsequent internet publications.

References

[1] B. Khan, R. Naseem, F. Muhammad, G. Abbas, and S. Kim, ''An empirical evaluation of machine learning techniques for chronic kidney disease prophecy,'' IEEE Access, vol. 8, pp. 55012–55022, 2020.

Jack Sparrow Publishers © 2024, IJRDES, All Rights Reserved
www.jacksparrowpublishers.com

ISSN: 2582 - 4201(Online) , http://www.ijrdes.com/, IJRDES, Vol. 6, Issue 3, 2024 , https://doi.org/10.63328/IJRDES-V6RI3P6

[2] S. Vallabhajosyula, L. Ya'Qoub, V. Kumar, D. Verghese, A. V. Subramaniam, S. H. Patlolla, V. K. Desai, P. R. Sundaragiri, W. Cheungpasitporn, A. J. Deshmukh, K. Kashani, and G. W. Barsness, ''Contemporary national outcomes of acute myocardial infarctioncardiogenic shock in patients with prior chronic kidney disease and end-stage renal disease,'' J. Clin. Med., vol. 9, no. 11, p. 3702, Nov. 2020.

[3] M. Almasoud and T. E. Ward, ''Detection of chronic kidney disease using machine learning algorithms with least number of predictors,'' Int. J. Adv. Comput. Sci. Appl., vol. 10, no. 8, pp. 1–9, 2019.

[4] D. Jain and V. Singh, ''A novel hybrid approach for chronic disease classification,'' Int. J. Healthcare Inf. Syst. Informat., vol. 15, no. 1, pp. 1–19, Jan. 2020, doi: 10.4018/IJHISI.2020010101. D. Jain and V. Singh, ''A two-phase hybrid approach using feature selection and adaptive SVM for chronic disease classification,'' Int. J. Comput. Appl., vol. 43, no. 6, pp. 524–536, Jul. 2021, doi: 10.1080/1206212X.2019.1577534

[5] R. Parthiban, S. Usharani, D. Saravanan, D. Jayakumar, U. Palani, D. StalinDavid, and D. Raghuraman, ''Prognosis of chronic kidney disease (CKD) using hybrid filter wrapper embedded feature selection method,'' Eur. J. Mol. Clin. Med., vol. 7, no. 9, pp. 2511–2530, 2021

[6] M. Schultz and M. Tropmann-Frick, ''Autoencoder neural networks versus external auditors: Detecting unusual journal entries in financial statement audits,'' in Proc. HICSS, 2020, doi: 10.24251/hicss.2020.666.

[7] G. A. Susto, A. Beghi, and S. McLoone, ''Anomaly detection through online isolation forest: An application to plasma etching,'' in Proc. 28th Annu. SEMI Adv. Semiconductor Manuf. Conf. (ASMC), May 2017, pp. 89–94.

Jack Sparrow Publishers © 2024, IJRDES, All Rights Reserved
www.jacksparrowpublishers.com